



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------|-------------|----------------------|---------------------|------------------|
| 09/824,126 | 04/02/2001 | James Gordon McLean | RPS9 2000 0118 | 9221 |

7590 04/22/2004

IBM Corporation
Personal and Printing Systems Group
Dept. 9CCA/Bldg. 002-2
P. O. Box 12195
Research Triangle Park, NC 27709

EXAMINER

LEWIS, MICHAEL A

| ART UNIT | PAPER NUMBER |
|----------|--------------|
|----------|--------------|

2655

DATE MAILED: 04/22/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/824,126

Applicant(s)

MCLEAN ET AL.

Examiner

Lewis A Michael

Art Unit

2655

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-37 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-37 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. ____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date 04.
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. ____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: ____.

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

2. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

3. Claims 1,6,17 & 22, are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) and further in view of Warnock et al (U.S. 6151576).

Regarding claims 1 & 17, Angell et al. disclose a method/computer medium for collaborative speech recognition in a network, comprising the steps of (Fig. 1):

- (a) Capturing speech as a plurality of audio streams by a plurality of capturing devices (Fig 1 (102)).

(b) Producing a plurality of text streams from the best quality audio stream by at least one recognition device (Col 3, Line 49 –53).

Angell et al do not disclose determining the best recognized text stream from the plurality of text streams. However, Warnock et al. teach the use of text stream reliability measure [choosing the best text stream] at the output of a speech recognition system (Col 4, Lines 10 – 20). In speech-to-text conversions it is beneficial to know the confidence level of speech recognition engine output so that the best text stream can be chosen.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify Angell et al. with the use of a text stream confidence measure as taught by Warnock et al. since it would have provided a more useful speech to text conversion system.

Regarding claims 6 & 22, the combination of Angell et al. and Warnock teaches the use of a database for storing the best recognized text stream (Fig 1 (110)).

4. Claims 2 & 18, are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) in view of Warnock et al (U.S. 6151576) and further in view of Tai et al (U.S. Patent 6618704).

Regarding claim 2 & 18, the modified Angell et al. do not disclose capturing step (a) further comprising: (a1) determining a best quality audio stream from a plurality of audio streams. However, Tai et al. teach the use of several audio sensors [claimed capturing devices] that uses an audio selector [arbitration device] that calculates the best audio source in order to determine the choice of camera. This technique is used to focus on speakers for applications such as video conferencing where a camera is automatically trained on a speaker among a plurality of speakers.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. by choosing the best audio stream as taught by Tai et al. since it would have enhanced the results of the speech recognition application.

5. Claims 3 & 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) in view of Warnock et al (U.S. 6151576) in view of Tai et al (U.S. Patent 6618704) and in further view of Perez-Mendez et al (U.S 5754978).

Regarding claim 3 & 19, the modified Angell et al. do not disclose the step further comprising: (a2) routing the best quality audio stream to a plurality of recognition

devices. However, Perez-Mendez teaches the use of routing speech to different speech recognizers that have slight differences (Fig 7). The text are then rejected or accepted based on the agreement in a comparator. This assures the most accurate result from the recognizer.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. by routing speech to several recognizers as taught by Perez-Mendez et al. in order to assure the most accurate result form the recognizer.

6. Claims 5 & 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) in view of Warnock et al (U.S. 6151576) and in further view of Perez-Mendez et al (U.S 5754978).

Regarding claims 5 & 21, the modified Angell et al disclose the determining step (d) comprises:

(d3) Correcting the interim best-recognized text stream to obtain the best-recognized text stream (Col 4, Lines15 – 26).

The modified Angell et al do not disclose (d2) Determining an interim best-recognized text stream (Col 4, Lines 5 – 14) and (d1) assessing agreement between the plurality of text streams. However, Perez-Mendez teaches

comparing taking the intermediate output of several speech recognizers and using a comparator to accept or reject their output (Fig 7; Abstract). The agreement/comparison method assures the best result from multiple text streams outputted by the recognizers.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. with an agreement/comparison among text streams and a choice of the best text stream as taught by Perez-Mendez et al. since an agreement/comparison method assures improved speech recognition performance.

7. Claims 7,8, 9, 23, 24, 25, 26 & 27, are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) in view of Warnock et al (U.S. 6151576) and in further view of Wymore (U.S. Patent 6631348).

Regarding claims 7,8, 9, 23, 24, 25, 26 & 27, the modified Angell et al do not disclose a capturing device where the capturing and recognition device are the same or that the capturing device is comprised speech recognition technology. However, Wymore discloses a capturing device that has the ability to recognize speech that includes speech recognition technology (Fig. 2A). The capturing device with speech recognition has the ability to adapt to ambient surroundings such as noise, speech level, etc.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. with the use of a capturing device which includes speech recognition technology as taught by Wymore et al. since it would made the speech recognition device more adaptable to different types of noise profiles.

8. Claims 10,28 & 35, are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) in view of Tai et al (U.S. Patent 6618704) and in further view of Perez-Mendez et al (U.S Patent 5754978).

Regarding claims 10,28 & 35, Angell et al. disclose a method/computer medium for collaborative speech recognition in a network, comprising the steps of (Fig. 1):

- (a) Capturing speech as a plurality of audio streams by a plurality of capturing devices (Fig 1 (102)).
- (b) Producing a plurality of text streams from the best quality audio stream by at least one recognition device (Col 3, Line 49 –53).

Angell et al. do not explicitly disclose determining the best quality audio stream from the plurality of audio streams. . However, Tai et al. teach the use of several audio sensors [claimed capturing devices] that uses an audio selector [arbitration device] that calculates the best audio source in order to determine the choice of

camera. This technique is used to focus on speakers for applications such as video conferencing where a camera is automatically trained on a speaker among a plurality of speakers.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. by choosing the best audio stream as taught by Tai et al. since it would have enhanced the results of the speech recognition application.

The modified Angell et al. do not disclose determining the best recognized text stream from the plurality of text streams. However, Perez-Mendez et al. teaches the use of routing speech to different speech recognizers that have slight differences in their configuration (Fig 7). The text are then rejected or accepted based on the agreement in a comparator. This assures the most accurate result from the recognizer.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. by routing speech to several recognizers as taught by Perez-Mendez et al. would have given the best agreement on the text streams resulting in enhanced speech recognition results.

Regarding claims 13,31 & 37, the modified Angell et al. discloses a database in which text streams are the best recognized text stream are stored (Fig 1 (110)).

9. Claims 12,14,15,16, 30, 32, 33, 34 & 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Angell et al (US 6513003) in view of Tai et al (U.S. 6618704) in view Perez-Mendez et al (U.S 5754978) and further in view of Wymore (U.S. Patent 6631348).

Regarding claims 12,30 & 36, the modified Angell et al disclose the determining step (d) comprises:

(d2) Determining an interim best-recognized text stream (Col 4, Lines 5 – 14).

(d3) Correcting the interim best-recognized text stream to obtain the best-recognized text stream (Col 4, Lines15 – 26).

The modified Angell et al do not disclose (d1) assessing agreement between the plurality of text streams. However, Wymore teaches comparing an input utterance [text stream] against stored trained utterances [text stream] to determine best ambient noise setting configuration of an input device (Abstract). The ability to set the noise profile will affect the accuracy of a speech recognition application.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. with accessing the text streams as taught by Wymore et al. since it would have enhanced the audio capture adaptive process resulting in improved speech recognition system

Regarding claims 14,15,16,32,33 & 34, the modified Angell et al do not disclose a capturing device is a device that recognition capability or comprises speech recognition technology. However, Wymore discloses a capturing device that has the ability to recognize speech or speech recognition technology (Fig. 2A). A capturing device with speech recognition has the ability to adapt to ambient surroundings such as noise, speech level, etc.

Therefore it would have been obvious to one of ordinary skill at the time of the invention to modify the modified Angell et al. with the use of a capturing device with speech recognition technology as taught by Wymore et al. in order to set noise profile since the adaptive noise feature would have enhanced the audio capture process resulting in improved speech recognition system.

Conclusion

1. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- | | | |
|------|-----------------|-----------------------|
| i. | Kanevsky et al. | U.S. Patent (6618704) |
| ii. | Witteman et al. | U.S. Patent (6243676) |
| iii. | Andersen et al. | U.S. Patent (6704707) |
| iv. | Warnock et al. | U.S. Patent (6151576) |

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael A. Lewis whose telephone number is 703 305-8730. The examiner can normally be reached on Monday through Friday, 8:30 am – 5 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on (703) 305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Lewis A Michael
Examiner
Art Unit 2655

Mal

3/16/2004



DORIS H. TO
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600